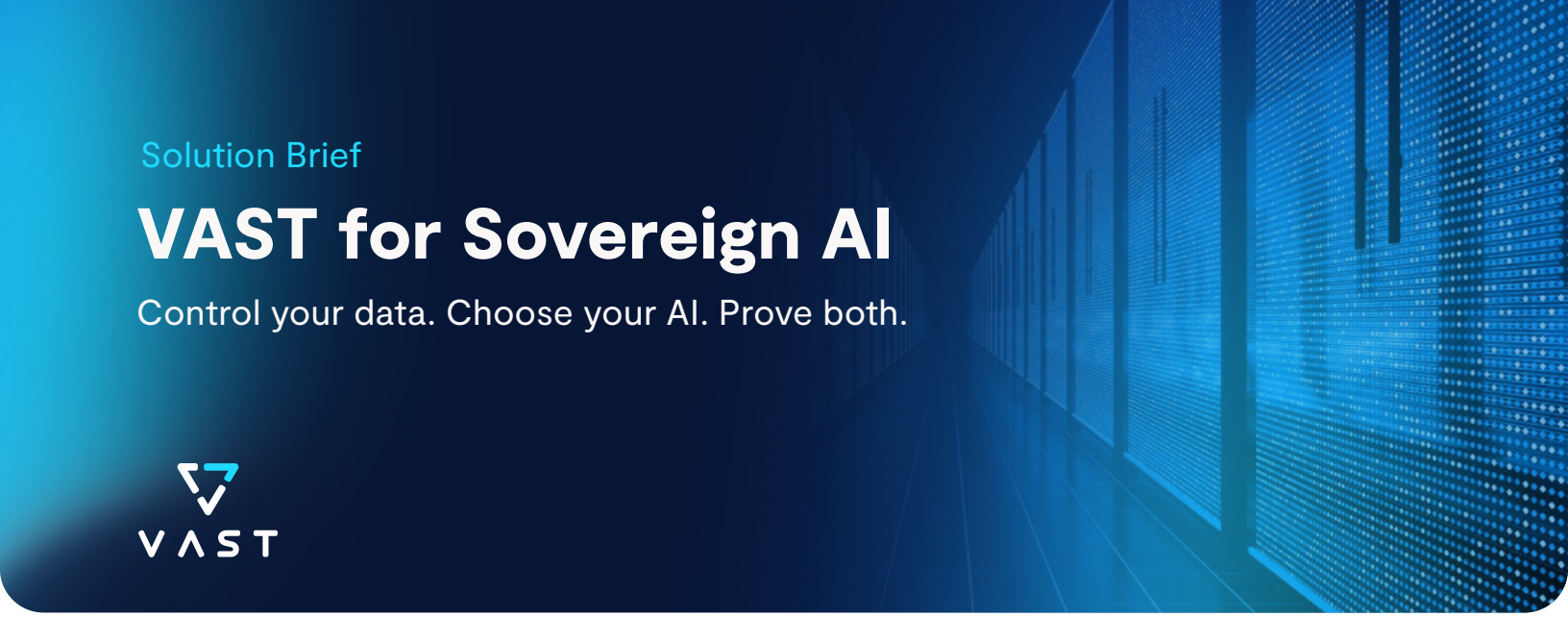
Solution Brief

VAST for Sovereign AI

Control your data. Choose your AI. Prove both.



Sovereign AI requires more than locating data in an approved country. Governments, regulated enterprises and critical infrastructure operators must control who can access sensitive information, how AI services use it, where derived data appears and what evidence survives an audit or incident.

Global AI, Locally Governed

AI creates new control surfaces: prompts, retrieved context, embeddings, vector indexes, checkpoints, inference logs, cache, model outputs, and agent actions. A source dataset may remain in-country while copies, metadata, or derived context cross an unapproved boundary. The task is to make AI-ready data accessible without multiplying unmanaged copies, policy systems, or supplier dependencies.

The VAST AI Operating System (AI OS) combines high-performance data access, shared metadata and AI data services to support an operator-defined sovereign deployment. VAST supplies infrastructure-level mechanisms; the customer defines jurisdictional policy, administrator authority, service boundaries, and acceptable dependencies.

Three Tests for a Sovereign Deployment

- **Territory:** Where data, replicas, metadata, and operational information reside.
- **Authority:** Who controls administration, access, policy, and keys.
- **Autonomy:** Whether services can continue, recover, and evolve under the operator's chosen dependencies. The design must also generate evidence of these controls.

One Data Foundation, Multiple Access Patterns

The VAST DataStore stores unstructured and structured data. Within VAST DataStore, the Element Store maintains protocol-independent data elements, while the protocol layer provides multiprotocol access to those elements. This unified approach extends to access control. VAST stores file and directory permissions internally in a protocol-independent format and translates them according to the security model configured for each view, helping maintain consistent authorisation across NFS, SMB, and S3.

Benefits

Unified Data Foundation

Bring file, object, block, table, metadata, and AI data services into one architecture to reduce avoidable copies and unify governance surfaces.

Governance from Flash to Vectors

Apply shared metadata and protocol-independent access controls across supported file, object, and AI retrieval paths; validate exact enforcement semantics.

Sovereign Federation and GPU Access

Use DataSpace to connect VAST environments and make datasets available to approved compute locations.

Protect AI Execution

Use DataEnclave confidential-computing patterns, attestation, and independent key domains to support AI on sensitive data without exposing both parties' assets.

Why This Matters for Sovereign AI

- Shared metadata and access models reduce discrepancies between file, object, SQL, and AI retrieval services when implemented within the supported configuration.
- Native NFS, SMB, S3, NVMe/TCP, and data-service interfaces support existing applications and new AI pipelines without requiring every workload to use a separate proprietary data store.
- A unified foundation reduces the operational burden of duplicate datasets, vector indexes, transfer pipelines, and audit collection points; it does not eliminate all copies or guarantee end-to-end lineage by itself.
- VAST's scale-out performance supports sustained GPU data access, helping operators avoid a false choice between tighter controls and AI throughput. Specific performance claims must be benchmarked in the target environment.

How VAST Contributes to Sovereign Control

Data Residency and Federation: VAST DataSpace

Federate namespaces and synchronise or replicate data between supported VAST clusters. Operators define which sites are permitted and how access, movement, and replication are configured. Federation must be accompanied by explicit residency, telemetry, support, and recovery policies

Tenant Isolation and Access: VAST AI OS

Support multi-tenant infrastructure with access permissions, separate security contexts, and encryption options. This matters when ministries, regulated organisations, research teams, or commercial AI customers share capacity but cannot share a trust boundary. Test administration, cross-protocol mappings, and tenant escape paths.

Confidential AI: VAST DataEnclave

DataEnclave introduces hardware-isolated execution and attestation-gated key release for selected confidential AI workloads. Customer data and model-provider intellectual property can remain protected in separate trust domains, including supported disconnected operating patterns. External model entitlement and hardware roots of trust remain dependencies.

Policy, Observability, and Recovery

Combine data access controls, audit capabilities, snapshots, recovery mechanisms, and operational logging to help demonstrate how selected policies operate. A complete compliance record can also require identity-provider, orchestration, model, application, network, and SOC logs. VAST alone does not reconstruct every AI action.

Control Model: Technology to Evidence

Customer Requirement	VAST Contribution	Assurance to Test
Keep sensitive data in defined sites	DataSpace placement and replication design	Replication, telemetry, diagnostics, support paths
Restrict access across AI pipelines	Identity integration and data-layer permissions	File-to-SQL/vector permission mapping
Protect sensitive inference	DataEnclave attestation and key domains	Attestation chain, revocation, and uptime dependency
Recover or switch under stress	Snapshots, open access protocols, federation	Restore exercises, export time, application dependencies

Reference Deployment: Controlled Data, Approved AI Compute

Consider a regulated organisation with sensitive records in customer-controlled infrastructure and a requirement to use external GPU capacity or proprietary AI models. The operator first classifies workloads and data. It then selects an approved compute environment, establishes identity and network trust, sets data location and replication policy, controls encryption keys, and defines logging and support-access rules.

Implementation Pattern

1. Keep authoritative datasets and metadata in designated VAST environments, with explicit controls over replication, backup copies, and diagnostic information.
2. Expose supported NFS, S3, SQL, or vector access paths to authorised workloads. Validate which permissions propagate and what each service logs.
3. Use DataSpace where federated access or managed transfer is appropriate; do not assume a global namespace means data never crosses borders.
4. Isolate tenants and define administrative roles, credentials, key-management responsibilities and approval workflows for models and agents.
5. Where confidential model execution is required, evaluate DataEnclave attestation and customer/model-provider key boundaries; plan for model-entitlement and hardware-trust dependencies.
6. Prove the design with access-denial tests, residency checks, audit exports, recovery drills, and documented supplier support procedures.

Build Your Sovereign AI Foundation with VAST

Discover how VAST AI OS, DataStore, DataSpace, and DataEnclave can help you design AI infrastructure around your organisation's requirements for data residency, access control, workload isolation, and operational flexibility.

One Data Fabric

Sovereign AI workloads run where compute capacity is available while maintaining control of data and security policies.

Maximum Utilization

GPUs can access and load models from VAST storage with performance consistent with local NVMe devices.

Consistent Policies

VAST DataSpace enforces a single trust zone across environments. Data governance remains consistent everywhere.

About VAST Data

VAST Data is the AI Operating System company – powering the next generation of intelligent systems with a unified software infrastructure stack that was purpose-built to unlock the full potential of AI. The VAST AI OS consolidates foundational data and compute services and agentic execution into one scalable platform, enabling organizations to deploy and facilitate communication between AI agents, reason over real-time data, and automate complex workflows at global scale. Built on VAST's breakthrough DASE architecture – the world's first true parallel distributed system architecture that eliminates tradeoffs between performance, scale, simplicity, and resilience – VAST has transformed its modern infrastructure into a global fabric for reasoning AI. Learn more at vastdata.com and follow VAST Data on LinkedIn, YouTube and X.

©2026 VAST Data, Inc. All rights reserved. All trademarks belong to their respective owners.

11/26 I 16619